



3D surface reconstruction and automatic camera calibration a Bayesian approach

March 2004

André Jalobeanu

*Esfandiar Bandari, Peter Cheeseman,
Frank Kuehnel, John Stütz, Doron Tal*

Bayesian Vision Group

NASA Ames Research Center - Moffett Field, CA - USA

Summary

- Goal?
- Proposed vs. classical methods
- Bayesian approach and graphical model
- The surface model*
- Observation parameters
- Forward problem : rendering*
- Inverse problem : optimization
- Preliminary results
- Applications
- Conclusion
- Extensions and future work

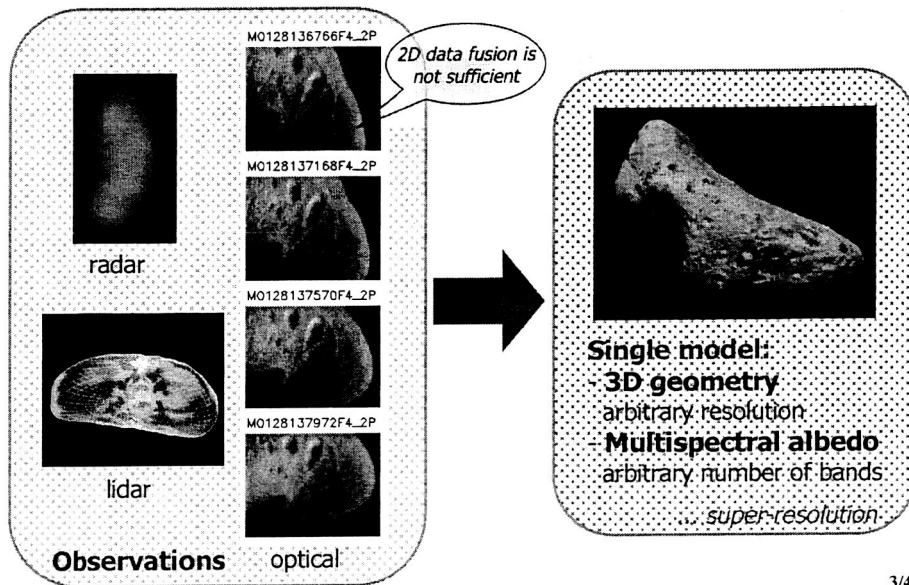
* *my contributions*

2/45



Ames Research Center

Long term goal: 3D data fusion



Ames Research Center

Case 1 : infer the geometry



Example - Asteroid 433 Eros (NEAR mission)

- N optical images; very large number $N > 10000$
- Light direction $\sim$ known
- Camera parameters $\sim$ known
- Albedo $\sim$ constant



N calibrated images

Geometry



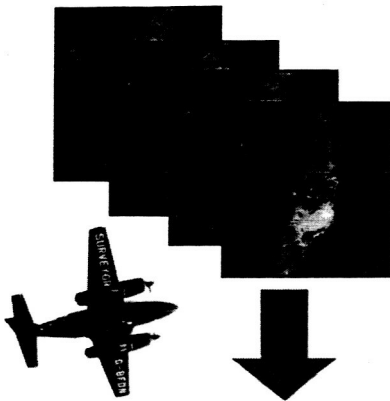
4/45



Case 2 : infer topography and scalar albedo



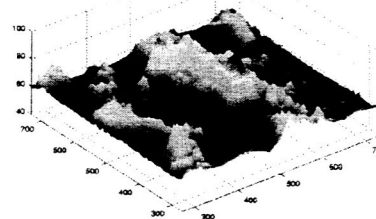
N aerial or satellite images



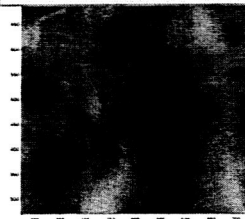
Cameras
Light sources



Elevation (height field)



Albedo (panchromatic)



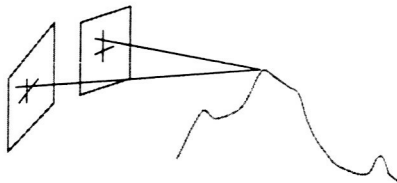
5/45



3D reconstruction: existing methods I



Shape from Stereo [Zhang et al. 94]



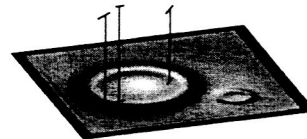
- The 2D projection of a surface point into the left and right cameras can be re-projected to give a point in the 3D space

Drawbacks :

- Relies on finding point **matches** in both images
- The **density** of the recovered surface points is data dependent

Shape from Shading [Horn & Brooks 89]

- Image gradients are related directly to surface gradients.
- Integrating from known boundary condition gives heights.



Drawbacks :

- Density of points is fixed at the **image resolution**
- **Assumptions:** Lambertian reflection and constant albedo

6/45

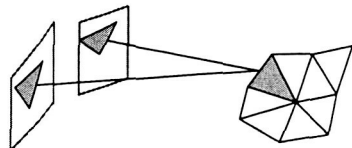


Ames Research Center

3D reconstruction: existing methods II



Generalized stereo [Fua & Leclerc 94]



- Use textured triangles on a dense 3D mesh as stereo features
- Optimize a cost function including a data term and a smoothness penalty term
- Combines stereo and shading information

Drawbacks :

- Not a Bayesian method: difficult to infer prior model parameters, and the data term is not natural

Other methods

- [Morris & Kanade 03]
Keep vertices fixed, search for good triangulations
- [Isidoro & Sclaroff 02, Vogiatzis et al. 03]
Stochastic algorithm
Iteratively shrink and deform a mesh

Drawbacks :

- Stochastic: slow, complex
- Others: sensitive to local optima

7/45



Ames Research Center

Advantages of the proposed approach



Existing methods are not compatible with our **long-term goal**:

dense reconstruction of both **geometry** and **reflectance**,
taking advantage of **all the information** in the **N** images

The proposed approach :

- Enables us to arbitrarily choose the surface **density**
(e.g. according to the amount of available data)
- Can use any **reflectance model**,
from the simple Lambertian BRDF to complex realistic BRDFs
- Can use data from **any sensor**, whatever the pixel size;
in principle we should be able to use non-optical data (e.g. altimetry)
- Should be able to easily integrate **new images**
once a initial estimate has been computed

8/45

The Bayesian approach

posterior density

$$P(\text{surface}, \text{cameras} | \text{images}) \propto$$

$$P(\text{surface}) \times \prod_i P(\text{camera}_i) \times \prod_i P(\text{image}_i | \text{surface}, \text{camera}_i)$$

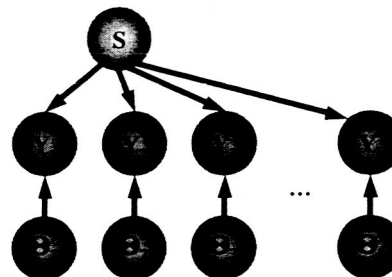
prior density likelihood

Surface model <i>Gaussian fractal model</i>	Prior camera pose <i>Gaussian model</i>	Image formation model <i>Rendering, blur, noise</i>
---	---	---

9/45

Graphical model (generative model)

- **Surface**
 - Mean: 3D object
 - Uncertainty on the object
- **Observations**
 - Mean: Rendering * Blur
 - Uncertainty from noise
- **Observation parameters**
 - Mean: calibration
 - Uncertainty on camera pose



Graphical model :
 Relationship between random variables
 (prior and conditional densities)

10/45



Surface model $P(S)$

Geometry: dense triangular mesh: vertices + neighborhood

Reflectance: one albedo per vertex

Topology: arbitrary – DEM (flat), planet/asteroid (spherical)
Subdivided mesh, has the topology of the initial mesh



Observation parameter density $P(\theta)$

Camera pose: position, orientatio

Camera physics : PSF, noise variance

Light source: orientation, ambient and direct intensity

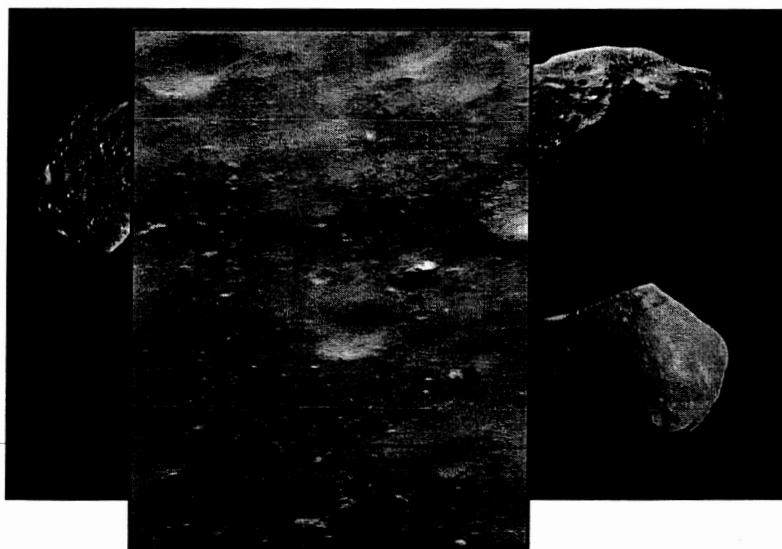


Image formation model $P(Y | I(S, \theta))$

Rendering: realistic synthetic image, $I(S, \theta)$ non-linear w.r.t. S, θ

Degradation: blur by PSF, Gaussian noise

11/45



12/45

What tool to use to analyze and model fractal surfaces?

Check the **statistical self-similarity**:

first compute the **power spectrum** of the object to analyze.

- Height fields on planar surfaces: *2D Fourier Transform*
- Height fields on spherical surfaces: *Spherical Harmonics*
- **Arbitrary surfaces** : ?
irregular sampling, arbitrary topology



Wavelets on subdivided meshes

Spectrum: feature size = fct(scale)

13/45

Topological support

Set of sites (vertices) + neighborhood system

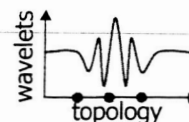
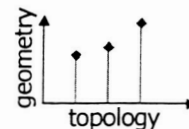
Support regularity = neighborhood regularity

Semi-regular mesh: 5 or 6 neighbors

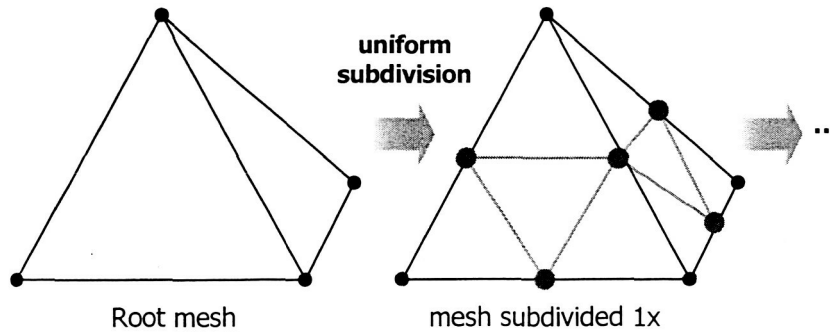
Geometry

3D point for each topological site

*Objects can have an irregular geometry,
but the wavelets are defined on a semi-regular
topological support.*



14/45

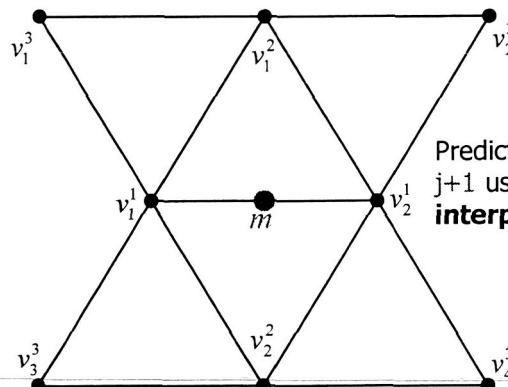


Creation of a new topological vertex at the midpoint of each edge
Each triangle is replaced by 4 smaller triangles

→ *ideal framework for a multiresolution analysis*

15/45

Geometric **Subdivision**: new vertex creation using a **prediction rule**



Prediction of a new vertex at level $j+1$ using 8 neighbors at level j :
interpolation scheme

[Dyn et al. 90,
Sweldens & Schroeder 95]

$$m = \frac{1}{2}(v_1^1 + v_2^1) + \frac{1}{8}(v_1^2 + v_2^2) - \frac{1}{16}(v_1^3 + v_2^3 + v_3^3 + v_4^3)$$

16/45



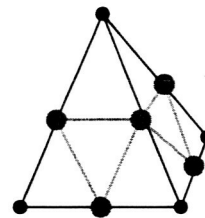
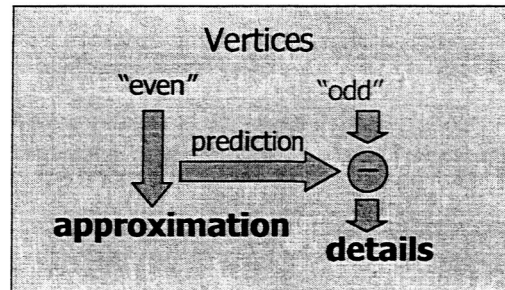
Ames Research Center

Wavelets on a triangular mesh



Wavelet coefficients at level $j+1$ =
vertices at level $j+1$ – prediction from level j

encode the **details** at level $j+1$



17/45



Ames Research Center

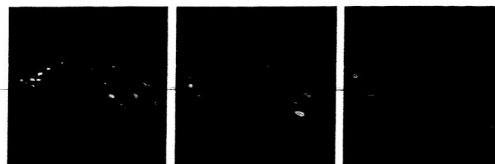
Multiresolution analysis



Approximations : $a_1 \dots a_n$
from coarse to fine, different versions of the same surface

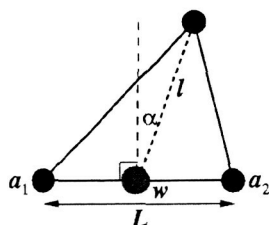
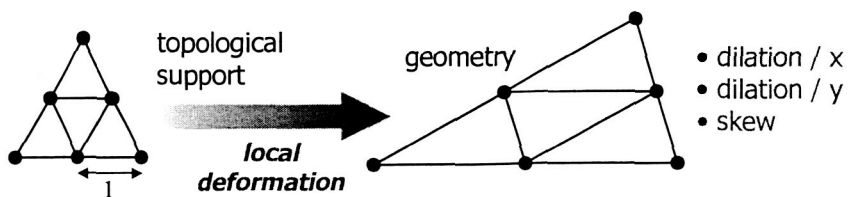


Details : $d_1 \dots d_n$
differences between two successive approximations



18/45

Defining a local scale



details = absolute geometric variations
(independent of the local mesh resolution)

→ we need to define a local scale estimate:

$$s = L \left(\frac{3}{4} \frac{L^2}{l^2} (\cos \alpha + \sin \alpha)^2 + 4 \right)^{-1/2}$$

19/45

Defining a local direction



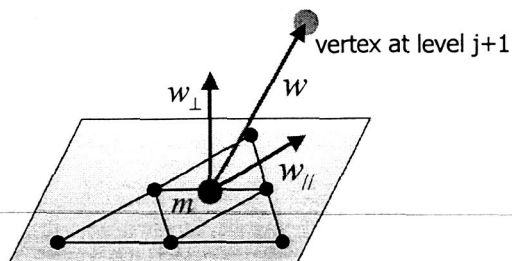
Wavelet details are 3D vectors

$$w = w_{\perp} + w_{\parallel}$$

normal + parallel decomposition

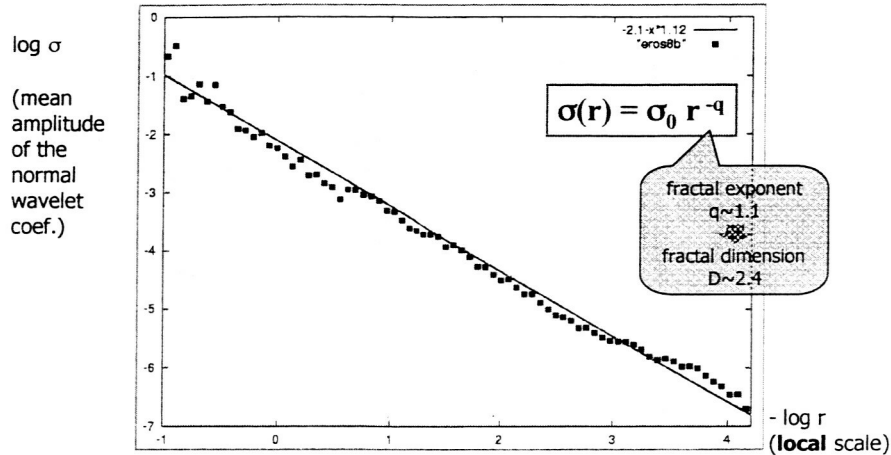
geometric detail

sampling irregularity



20/45

Amplitude spectrum of 433 Eros, data from the NEAR mission lidar



21/45

- Scale-invariant adaptive Gaussian model on $w_\perp$: **prior roughness** of the surface

$$P(\{w_\perp\}) \propto \prod_i \exp\left(-\lambda_i(s_i)^{-2q} |w_{i\perp}|^2\right)$$

geometric details =
Gaussian random variables

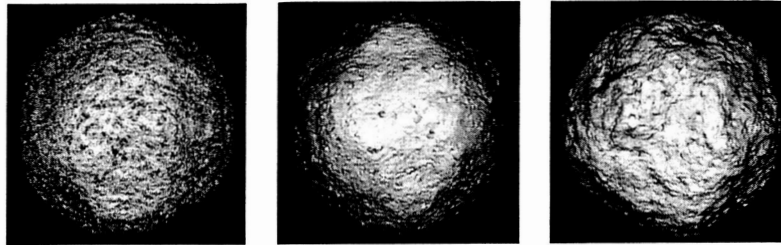
spatially adaptive
parameters

local scales

~ 3D analog of the **fractional Brownian motion** in 2D
(wavelet coefficients instead of Fourier coefficients)
Efficient description of the power spectrum of natural images

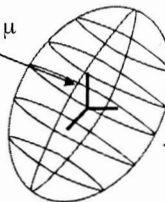
- Statistical model of $w_\perp$: mesh regularity prior density
(sampling regularity)

22/45



23/45

prior camera pose : mean μ



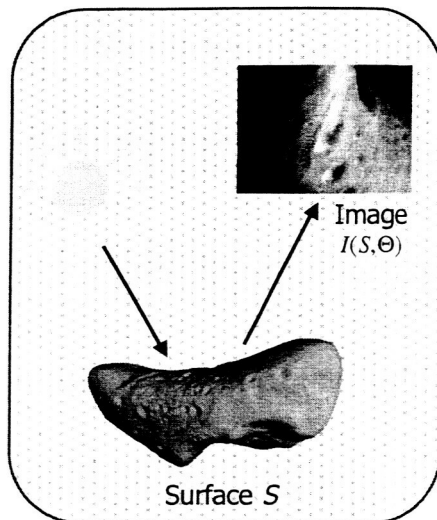
uncertainty : covariance Σ

Prior on the observation parameters (camera, light, PSF, noise) :

- Knowledge of the camera motion and orientation (rover: odometers; space probe: gyroscopes, etc.)
- Use the navigation parameters if available
- Sensors: orientation w.r.t. sun or stars
- Prior knowledge of the PSF and noise level



24/45



Rendering :

- known **surface**
- known **reflectance** } S
- known **light source**
- known **camera** param. (internal & external) } Θ



Compute the intensity for each pixel
... and its derivatives !

25/45

- **Sub-pixel** accuracy required for good reconstruction accuracy and super-resolution
- Compute the **derivatives** of the pixel intensity w.r.t. all the parameters (surface, camera, light) to perform optimization
- **Photometric accuracy**: ambient light, reflectance functions, etc.
- **Occlusions** (hidden surface removal)
- **Shadows** (remove surfaces hidden from the light source)



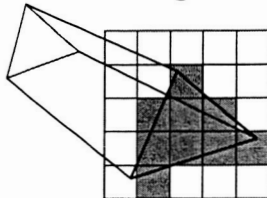
We need to work in the object space

Most of existing algorithms work in image space

26/45

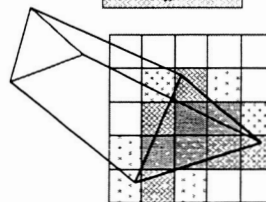
Image space or object space?

Image



- Very fast, real time (OpenGL, etc.)
- Image = non-continuous function of the surface S
- Aliasing
- Limited to big triangles

Object



- Sum of contributions / pixel

$$I_p \approx \sum_{\Delta} A_{\Delta}^p \Phi_{\Delta}$$
- Image = continuous fct. of S
- No aliasing (better sampling)
- Works with any triangle size
- Independence image/model

27/45

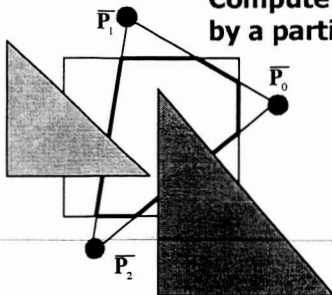
The rendering method

Core of the algorithm : **occlusion removal**

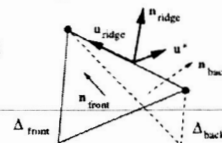
Hidden surfaces = occlusions / camera projection

Surfaces in shadow = occlusions / light source projection

Compute the contribution to a pixel by a partially hidden triangle



- Polygonal approach
- Recursive polygon-triangle subtractions
- Trick:
only on ridge lines



- At the end: polygon moments + derivatives

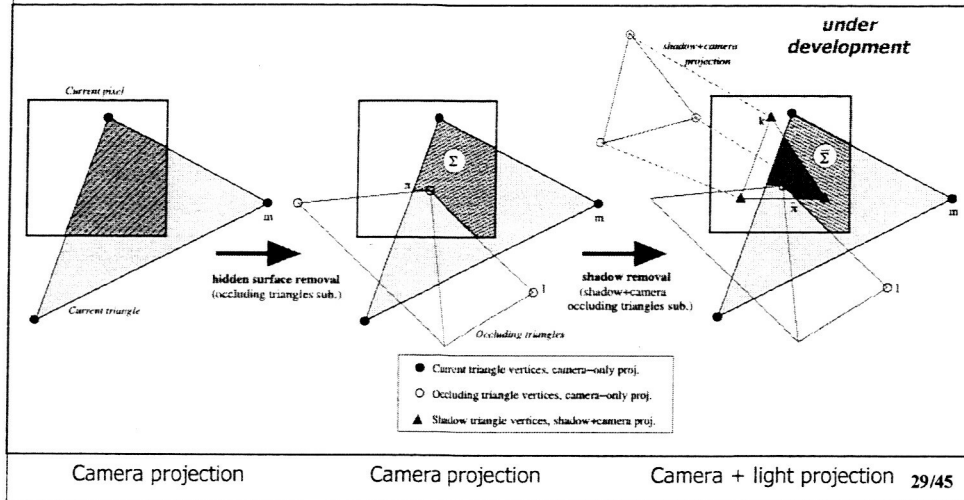
28/45

Rendering method - details

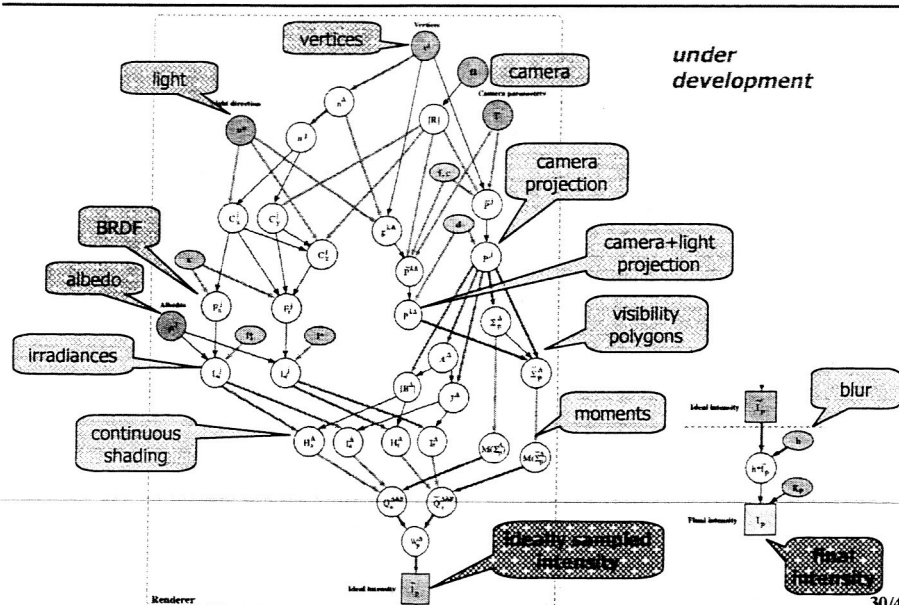
Intersection
triangle/pixel

Occluding triangles
subtraction

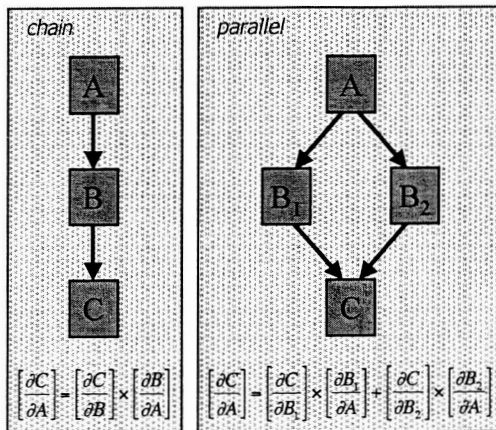
Shadowing triangles
subtraction



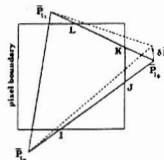
The rendering tree...



The chain rule :



Examples of derivatives :



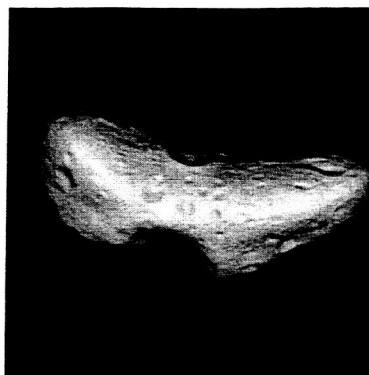
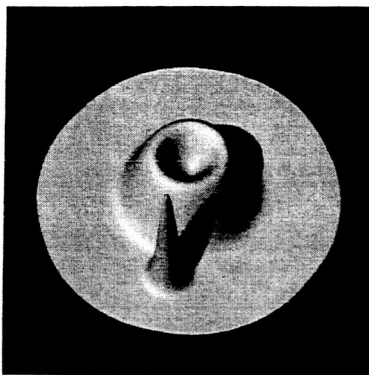
$\left[\frac{\partial P_j}{\partial v_j} \right]$	projected vertex	$\left[\frac{\partial n_j}{\partial v_j} \right]$	normal
$\left[\frac{\partial \pi_i}{\partial P_j} \right]$	2D polygon vertex	$\left[\frac{\partial L_j}{\partial n_j} \right]$	irradiance
$\left[\frac{\partial A}{\partial \pi_i} \right]$	pixel intensity		

31/45



32/45

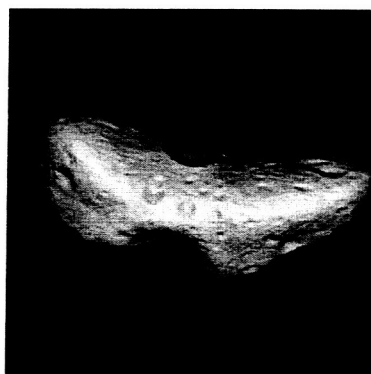
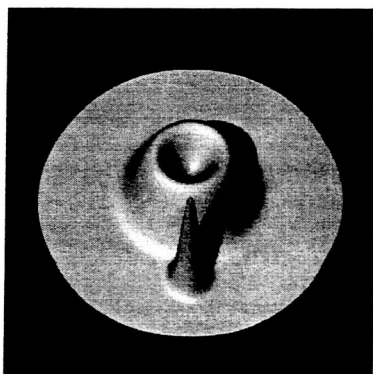
Rendering examples 1



Light source rotation

33/45

Rendering examples 2



Object rotation

34/45



Solving the inverse problem: optimization



Maximum A Posteriori (MAP) - iterative optimization

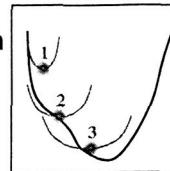
- **Linearize** the intensity (rendering):

$$I(S, \Theta) \approx I(S_0, \Theta_0) + \left[\frac{\partial I}{\partial S} \right] (S - S_0) + \left[\frac{\partial I}{\partial \Theta} \right] (\Theta - \Theta_0)$$

$\Rightarrow -\log P(S, \{\Theta\} | \{Y\})$ approx. by a **quadratic form**

$\rightarrow$ optimization of $S, \{\Theta\}$ using a *conjugate gradient*

- Result: used to initialize the next iteration
- Convergence: small variation of $S, \{\Theta\}$



- Optimization over S :
3D geometry and reflectance recovery
- Optimization over $\{\Theta\}$:
automatic camera and light calibration

35/45



Estimating the uncertainties

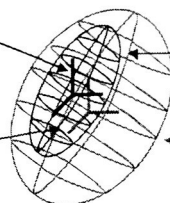


estimated pose : mean μ

uncertainty : covariance Σ

prior pose

prior uncertainty



- At the end : keep the **inverse covariance matrix** (related to the uncertainty), not only the estimated parameters $S, \{\Theta\}$
- Very useful to initialize a new estimation procedure, so that new data can be added
- **recursive refinement** of $S, \{\Theta\}$:
 - Update the 3D (process large amounts of data recursively instead of batch)
 - Refine the camera pose - Simultaneous Localization And Mapping (SLAM)

36/45



Ames Research Center

Preliminary surface reconstruction results Duckwater, Nevada



- Physical model built using USGS elevation model
- Hand-painted albedo



One of the 8 observed images

Images : CMOS camera
(10 bits, monochromatic)
Light source : sun

Goal: reconstruct DEM+albedo

Assumptions:
no shadows, no occlusions [old]

Preliminary estimation :

- sun direction (sun dial)
- camera param. (checkerboard)

37/45

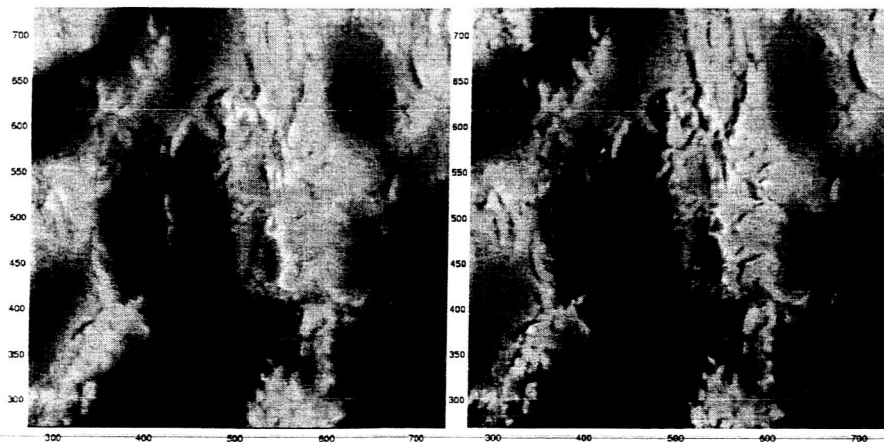


Ames Research Center

Preliminary surface reconstruction results Duckwater, Nevada



Geometry (height field)



Ground truth (USGS DEM)

Inferred DEM
(max error<10mm, RMS error<2mm)

38/45

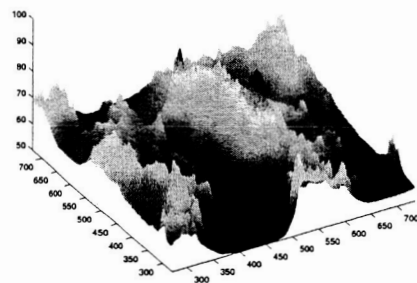


Ames Research Center

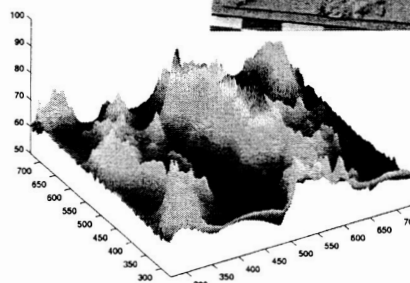
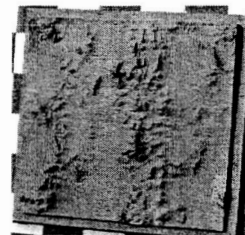
Preliminary surface reconstruction results Duckwater, Nevada



Special case: uniform albedo



Ground truth (USGS DEM)



Inferred DEM
(max error ~ 15mm)

39/45



Ames Research Center

Estimating the camera parameters

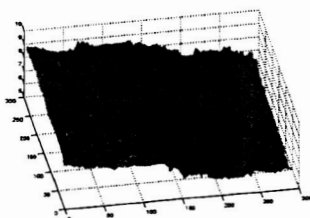


Camera pose estimation

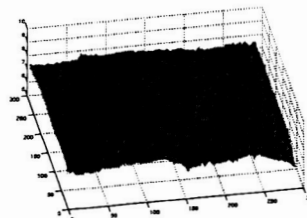
Method used : joint MAP = estimation of $(\{\Theta\}, S)$

Alternate optimizations w.r.t. surface and camera (sub-optimal)

Estimation error



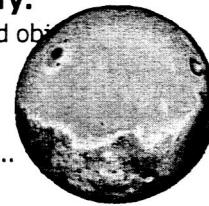
Camera calibrated from points



Camera calibrated using MAP

40/45

- **3D object reconstruction from multiple images:**
 - Asteroids (albedo ~uniform, spherical topology)
 - Planetary surfaces (variable albedo, planar/spherical topology)
 - Rover situation: MER mission (variable albedo, overhangs)
- **Simultaneous Localization And Mapping (SLAM):**
 - Simplified terrain model: use 3D features instead of meshes
- **Space probe localization and 3D object recovery:**
 - approach and flyby: recursive trajectory estimation and object model refinement
- **Multi-sensor data fusion:**
 - Optical (multi- and hyperspectral), radar (SAR), lidar, ...
- **Fractal geometry and synthetic images:**
 - Generate realistic fractal surfaces
 - Compute photorealistic BRDFs for natural surfaces



41/45



- **Computer vision**
(3D model reconstruction from multiple observations):
inverse problem of rendering
- **Bayesian inference**
applied to this inverse problem:
everything is described by random variables
- 2D data fusion into a single 3D model
becomes a **parameter estimation** problem
- It can be solved by existing efficient
optimization techniques

42/45

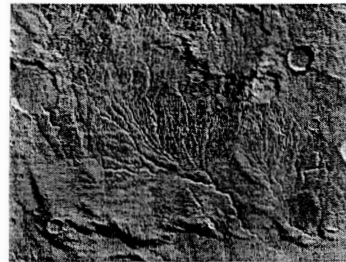
• Rendering:

- Take into account shadows and occlusions
- Visibility polygons (recursive subtraction)
- Derivative computation
- *Extensions*



• Surface modeling:

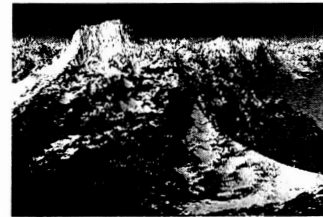
- ~new wavelet transform on surfaces of arbitrary topology:
- normal/parallel wavelet detail separation
- local scale computation
- Fractal model for natural surfaces



43/45

• Extensions (in progress):

- Accurate shadows
- Continuous shading (visibility polygon moments)
- Continuous PSF, take into account the blur
- complex BRDFs ($\neq$ Lambert)



• Future extensions:

- Multispectral
- Adaptive subdivision
- Adaptive BRDFs
- Secondary reflections
- Other types of camera (push-broom, etc.)
- Approximate methods



44/45



- **Model extensions:**

- Study real surfaces (Earth, Mars):
 - Reflectance/geometry interactions
 - Multi- and hyperspectral albedos



- **Inference method extensions:**

- Marginalization:***

- Simultaneous reconstruction and calibration
 - Separated geometry and albedo inference

- Bayesian model selection:***

- Infer the topology (overhangs, etc.)
 - Dynamic and adaptive subdivision

